MPEG-4, ИЛИ СТРУКТУРИЗАЦИЯ МУЛЬТИМЕДИА (часть 2)

Статья была опубликована в журнале PC Magazine, дополнена и переработана специально для журнала РЭТ.

(Окончание. Начало в РЭТ №6, 2003 г.)

Сергей Новосельцев (Москва)

В этом номере журнала мы продолжаем рассказ об информационной технологии грядущего цифрового мира – стандарте MPEG–4.

ПОТОКИ И УРОВНИ

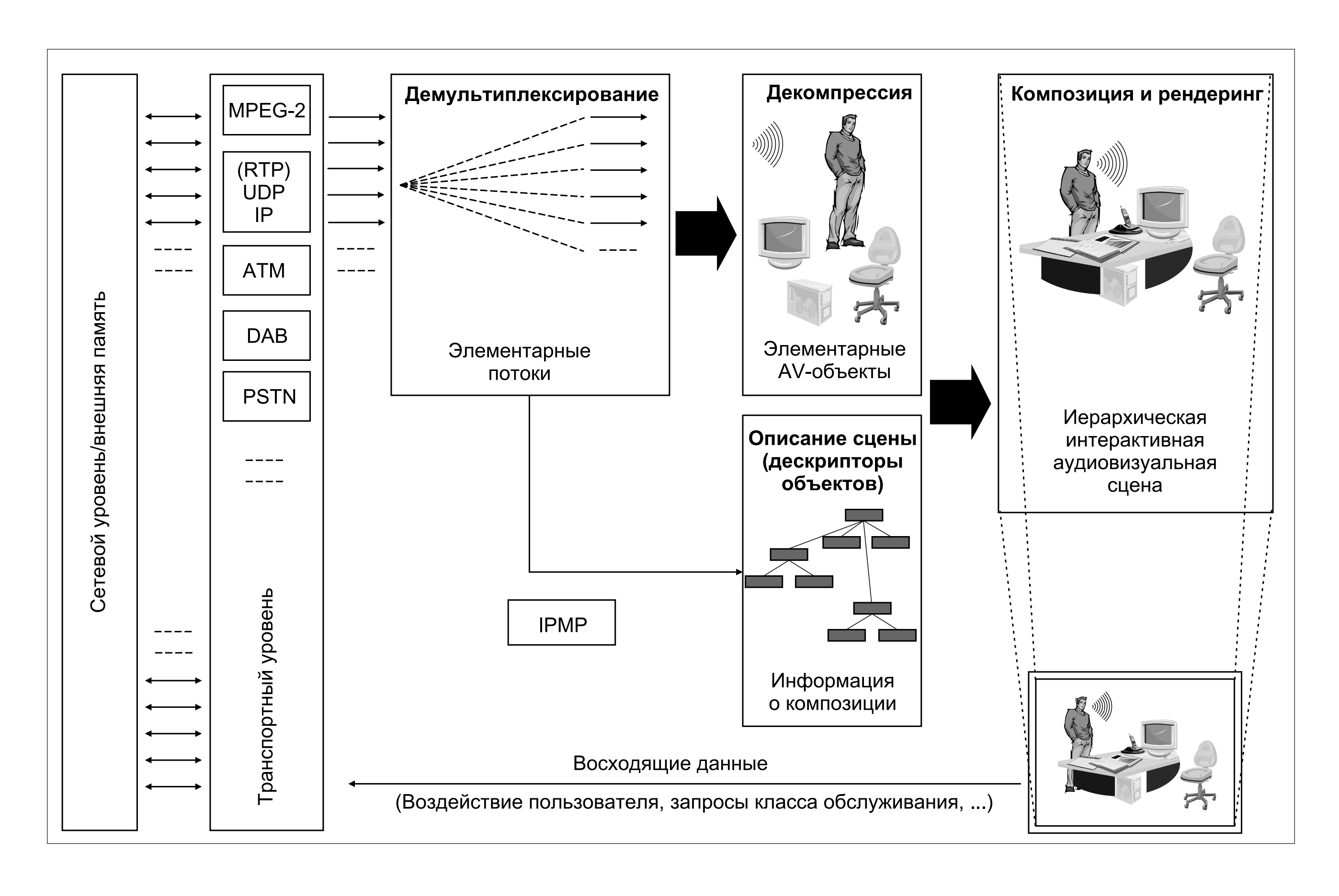
В «потоковой» части архитектура MPEG–4, очевидно, опирается на другой фундаментальный стандарт ISO – семиуровневую модель взаимодействия открытых систем. Напомним, эта модель выделяет семь независимых вложенных уровней (сверху вниз: прикладной, представительский, сеансовый, транспортный, сетевой, канальный, физический). Каждый из уровней на передающем конце общается с соответствующим уровнем на приемном, а для этого обращается к локальным службам соседнего нижнего уровня (который предоставляет ему для этого специальный интерфейс) – и далее все происходит «прозрачно» для него, все остальные нижние уровни системы от него скрыты. По мере продвижения по этой лестнице вниз содержательная информация, которой обмениваются пользователи или приложения – т.е. прикладные уровни, – обрастает служебными данными (которые «навешивает» каждый из уровней для взаимодействия со своим визави на приемном конце в соответствии с принятым между ними протоколом общения). Данные разбиваются на пакеты, тем или иным способом мультиплексируются, кодируются и передаются между узлами сети, а на приемном конце происходит обратный процесс – сообщение собирается, очищается от вспомогательных данных и восстанавливается до своего первоначального вида.

MPEG–4 действует на верхних уровнях модели ISO, начиная с сеансового. Для передачи потоков данных он обращается к службам транспортного уровня, которые обеспечивают приложениям инвариантность работы с различными системами и средами доставки – сетевыми (в том числе IP/UDP/RTP, ATM, сетями коммутации пакетов H.223), вещательными (кабельные и спутниковые системы, DTV, DAB), дисковыми (CD, DVD). (Что касается более низких уровней, собственно сетей и технологий передачи мультимедиа информации, то можно сослаться на весьма полный обзор, составленный Олегом Фоминовым [5].)

Для управления передачей потоковых данных в MPEG–4 предусмотрен специальный протокол сеансового уровня, называемый DMIF (Delivery Multimedia Integration Framework – среда интеграции доставки мультимедиа). Разработчики указывают на его сходство с FTP, подчеркивая при этом, что основное различие в том, что «FTP в ответ на запрос передает данные, а DMIF – указатели на то, где находятся (потоковые) данные». Службы уровня DMIF в декодере MPEG–4 устанавливают сеанс с «передающей стороной», затем выбираются нужные потоки, посылается запрос, в результате чего транспортный уровень устанавливает требуемые соединения, по которым будут поступать потоковые данные, и сообщает указатели на эти соединения. В итоге устанавливается прямой канал обмена данными между приложениями.

Службы DMIF доступны прикладному уровню с помощью интерфейса DAI (DMIF–Application Interface). Именно DAI маскирует для локальных приложений разницу между сетевыми, вещательными и локальными (например, с дисков CD/DVD) потоками, эмулируя при работе с вещательными и дисковыми источниками «удаленный DMIF» и «удаленное приложение». При этом допускается одновременная работа со всеми тремя типами источников и замена одного на другой.

Потоковые данные, которые относятся к медиаобъекту, могут поступать через один или несколько элементарных потоков (elementary streams, ES). Все необходимые характеристики этих потоков, как-то: требования к приемнику, данные о тайминге и об уровне обслуживания (Quality of Service, QoS), т.е. скорости, приоритете, допустимом уровне ошибок и максимальной задержке, содержатся в предусмотренном для каждого объекта дескрипторе объекта. Дескрипторы могут также содержать текстовую информацию об объекте. Дескрипторы объектов передаются в специальном элементарном потоке, что позволяет добавлять к сцене новые объекты или удалять ненужные динамически. Команды описания сцены и объектов в формате BIFS также составляют отдельный элементарный поток и могут быть модифицированы без изменения собственно медиаданных в других потоках. Таким образом можно выстраивать различные сценарии на базе одних и тех же медиапотоков. Это пригодится и для интерактивных применений с многовариантным развитием сюжета, и для подстройки сложности сцен и объектов под доступный уровень QoS: в обоих случаях можно заранее предусмотреть несколько командных BIFS–потоков, а в процессе передачи оперативно выбирать из них подходящий в качестве действующего сценария. Кроме того, облегчаются адаптация готовых произведений для новой среды доставки (например, CD–продуктов для WWW) или извлечение готовых объектов для использования в новых произведениях.



Упрощенная блок-схема декодера MPEG-4

Стандартом предусмотрено наличие отдельного, общего для всех типов потоковых данных уровня синхронизации, который для каждого элементарного потока определяет минимальную единицу доступа, или access unit (т.е. аудио- или видеокадр, команду описания сцены и т.п.), выстраивает для каждого объекта и для всей сцены временн'Ую базу и обеспечивает синхронизацию между ними. На синхроуровне элементарные потоки разбиваются на пакеты, к ним добавляется информация тайминга (time stamps) – чтобы на приемном конце декодер смог адекватно собрать и отобразить результирующий поток. Затем специальный подуровень мультиплексирования (FlexMux) определяет элементарные потоки с близкими требованиями к QoS и группирует их – для того, чтобы минимизировать число сетевых соединений, запрашиваемых от транспортного уровня. Сам транспортный уровень (как и более низкие) в стандарте не рассматривается, однако установлены методы защиты от ошибок, ресинхронизации и восстановления данных при сбоях в механизме доставки.

На приемном конце происходит декодирование потоков, выделение объектов и построение сцены. Особо подчеркнем, что, как и в случае MPEG-1 и MPEG-2, MPEG-4 не устанавливает правил процесса кодирования; не касается он и деталей реализации декодера, задавая лишь правила поведения некоего абстрактного устройства, а также синтаксис и семантику двоичных потоков, с которыми оно должно уметь работать. Для этого в MPEG-4 определена модель декодера – System Decoder Model. На практике допустимы всевозможные реализации декодеров MPEG-4, от отдельных специализированных терминалов до функций, встроенных в телевизор или приставку, от мобильных коммуникационных устройств вроде телефона до программных модулей в ПК, с разной степенью сложности (см. ниже).

Упрощенная блок-схема декодера MPEG-4 приведена на рисунке.

АВТОРСКИЕ ПРАВА

Одна из самых острых и трудно решаемых проблем цифрового мира, проблема, сегодня уже тормозящая его развитие куда сильнее, чем проблемы технологические, – это защита авторских прав. Стандарт же MPEG-4, претендуя на роль универсальной среды доставки контента, немедленно становится средоточием и болевой точкой этой проблемы. Понимая это, разработчики стандарта с самого начала привлекли представителей различных творческих профессий и отраслей, стремясь добиться единого подхода, определить синтаксис и набор средств идентификации и защиты прав интеллектуальной собственности (IPR) для MPEG-4. В результате был

выработан комплекс мер, известный под названием IPMP (Intellectual Property Management & Protection), подробное изложение которых выходит за рамки этой статьи (ссылки на соответствующие документы можно найти на сайте MPEG). Вкратце упомянем, что с дескриптором каждого объекта или потока может быть связан специальный блок данных (IPI, Intellectual Property Identifier), содержащий уникальный идентификатор в одной из принятых международных систем (ISAN, ISRC или др.), характеристику типа контента, а также имя обладателя прав или указатель на него (на них). Каждый декодер имеет блок (интерфейс) IPMP, который обрабатывает данные о защите. Стандартом предусмотрены также точки входа для шифрования/дешифрования информации. Выработанная система позволяет реализовать механизмы отслеживания авторских прав, автоматического отчисления авторских процентов, проведения аудита и расследований в случае предполагаемых нарушений, строить разные уровни защиты контента — по соображениям коммерческим, личностным, секретности и т.п. Однако эти «верхние» уровни защиты и управления правами не входят в стандарт и могут быть реализованы разработчиками приложений и/или держателями контента.

Естественно, и эти меры, особенно при наличии программно реализованных плееров MPEG-4, в принципе можно обойти, так что работа над проблемой авторских прав продолжается.

«ПРОФИЛИ» MPEG-4

Как видите, стандарт MPEG-4 в его полном виде — весьма разветвленный и многоплановый конгломерат, включающий множество механизмов и инструментов, так что его полная реализация может показаться задачей почти невыполнимой. Чтобы не допустить неконтролируемого размножения малосовместимых частичных реализаций, был установлен набор подмножеств, которые содержат ограниченные наборы инструментов и функций MPEG-4, существенных для тех или иных применений. Эти подмножества были названы «профилями» (Profile), они могут частично пересекаться, полностью включать в себя функциональность «младших» подмножеств или добавлять те или иные функции. Для большей гибкости и упрощения подбора вариантов профили были разбиты по категориям: девять визуальных (включающие, в свою очередь, профили для работы только с живым видео, только с анимацией и гибридные), четыре звуковых, три графических и четыре профиля описания сцены. Кроме того, в зависимости от доступной вычислительной мощности декодера, для каждого профиля установлены один или несколько уровней (Level — не путать с Layer модели ISO). Таким образом, при построении декодера MPEG-4 разработчик должен выбрать комбинацию профилей и уровней и после этого обязан реализовать описываемый ими набор функций в полном объеме. Потребитель же, прочтя в паспорте устройства или программы эту комбинацию, сразу понимает, что умеет, а чего не умеет данный декодер. Реализации, построенные на основе одинаковой комбинации, должны быть полностью совместимы друг с другом. Естественно, в предельном случае, выбрав комбинацию из всех старших профилей и уровней, мы получим полный набор функций MPEG-4. Список предлагаемых профилей можно найти в [1] (Версия 1) и [2] (Версии 1 и 2).

ВЕРСИЯ 2

В октябре 1998 г. был зафиксирован набор полностью готовых на тот момент функций и инструментов MPEG-4, этот набор был назван MPEG-4 Version 1 и передан на утверждение в качестве официальной спецификации стандарта. Все последующие доработки вошли в MPEG-4 Version 2. Версия 2 не заменяет функции Версии 1, а добавляет к ней новые возможности, сама же Версия 1 ревизии не подлежит. Декодеры, построенные по Версии 1, не устарели с выходом Версии 2, поскольку новые функции реализованы как дополнительный набор профилей.

Что же добавляет вторая версия стандарта? Список новых функций занимает не одну страницу, мы назовем лишь некоторые из наиболее впечатляющих.

- Разрешение многопользовательского присутствия в сцене и взаимодействия с контентом. Значит ли это, что в принципе можно залезть через сеть в чужой телевизор и все там передвинуть? или зайти в гости к соседу в виде аватара и сыграть с ним в Quake?
- Дальнейшее пополнение функций BIFS (BIFS++) и сближение его с VRML.
- Введение формата файла MPEG-4 (MP4) на базе формата файла QuickTime, что позволит хранить (локально или распределенно с URL-ссылками), копировать, редактировать, проигрывать на локальном устройстве и передавать полную презентацию в формате MPEG-4.
- MPEG-J позволяет загружать (в отдельном потоке) и запускать код Java на плеерах MPEG-4.
- Работа с полигональными 3D-моделями, поддержка уровней детализации модели (LOD).
- Вдобавок к анимации лица появляется анимация тела (разработано совместно с Humanoid Animation Working Group VRML).
- Стереоскопическое видео.
- Значительные усовершенствования на уровне DMIF, включая симметричные соединения отправитель—получатель вместо сервер—клиент, что позволит строить разговорные приложения и организовывать поиск в мультимедийных базах данных.
- и многое другое.

В результате число профилей в каждой из категорий как минимум удвоилось...

* * *

Несмотря на беглое и схематичное изложение, объем статьи не позволил рассказать обо всех хитрых свойствах MPEG-4 и оценить многочисленные удивительные возможности и перспективы, которые сулит его внедрение (хотя на это и трудно было рассчитывать — ведь даже краткий обзор стандарта, сделанный разработчиками, занимает более 60 страниц!). Но мы видели свою задачу в том, чтобы создать ощущение надвигающихся перемен и побудить к дальнейшим самостоятельным поискам — и потому отсылаем читателя к приведенной в конце статьи литературе, а также рекомендуем исследо-

вать сайт разработчиков MPEG по адресу http://mpeg.telecomitalia.com и сайт www.m4it.org. Справочный сайт www.mped.org уже несколько лет не обновлялся, но по-прежнему содержит много полезных ссылок; он может быть крайне интересен для тех, кто хочет разобраться с историей вопроса.

РАЗ, ДВА, ЧЕТЫРЕ, СЕМЬ, ИЛИ MPEG КАК ИНДИКАТОР ПРОГРЕССА

За первые несколько лет работы комитета MPEG у широкой аудитории успело сложиться представление о нем как об организации, целиком посвятившей себя вопросам сжатия медиаданных, – и это одна из причин, почему так поражает при первом знакомстве содержание стандарта MPEG-4. Но если разобраться, то даже само название MPEG – «Экспертная группа по движущимся изображениям» – показывает, что круг ее интересов значительно шире проблем сжатия. Просто на первом этапе «мультимедиа революции» именно сжатие имело решающее значение, и в MPEG уделили ему наибольшее внимание, добившись, отметим, беспрецедентно успешного результата в деле примирения подходов и интересов многомиллиардных корпораций и целых индустрий.

Сегодня происходит сближение (начинается интеграция) телевидения и Интернета, персональных компьютеров и развлекательных приставок и плееров, а медиаконтент, который потребитель получает из всех этих источников, становится не просто цифровым, но и все более интерактивным. И требуются новые стандарты, которые помогут разработчикам контента донести свои произведения до потребителя максимальным числом способов, а пользователям (зрителям) – получать со своего устройства доступ к медиаконтенту в любой его форме.

В целом можно сказать, что разработчики MPEG-4 собрали и обобщили многое из того, что было наработано за десять-пятнадцать лет в ранее мало пересекавшихся областях и технологиях (QuickTime и VRML, 3D-графика и интерактивная «персонажная» анимация по типу Macromedia Director, разработка видеоигр, видеокомпозитинг, телевещание, потоковые видео и звук), и сумели объединить все это в новое качество.

Насколько быстро MPEG-4 станет общепринятым, и станет ли? Когда вслед за появлением первых сетей с коммутацией пакетов ISO была впервые предложена семиуровневая модель, многие упрекали ее в избыточности и тяжеловесности, в «сложности в реализации», старались выкинуть «лишние» уровни или вообще обойтись без нее, предлагая альтернативные концепции. Однако постепенно модель была воспринята всеми и по существу легла в основу устройства всего созданного за эти годы связанно-цифрового мира – даже мыслить многие стали в категориях этих семи уровней. Подобным же образом в 1991 г., в весеннюю пору «цветения ста цветов» в области сжатия мультимедиаданных, критически воспринимались разработки группы MPEG, о проекте стандарта говорили как о «слишком сложном, чтобы быть реализованным», сетовали на высокий уровень требующейся аппаратной поддержки... Сегодня уже не только MPEG-1, но и MPEG-2 может быть реализован программно, на базе этих стандартов выстроились целые индустрии, стандарты MPEG отмечены наградами Emmy.

Поэтому, несмотря на кажущуюся (особенно для вещателей!) сложность MPEG-4, имеет смысл подробно разобраться с этим стандартом – он с очень большой вероятностью определит развитие компьютерных, вещательных и даже мобильных систем в ближайшие годы. Подчеркнем еще раз – разбираться придется не только программистам, для которых открывается новая интересная ниша (реализация сложной и многосторонней функциональности MPEG-4 для самых разных клиентских и серверных платформ), но и гуманитарно-творческим людям, авторам и продюсерам вещательных программ. Вслед за интерактивным контентом на дисках и в Сети интерактивными становятся телепрограммы – и навыки программирования и алгоритмического мышления очень пригодятся разработчикам этих новых программ. Давний термин «TV programming» – телевизионное программирование, – означавший искусство составления сетки вещания, приобретает новое, теперь вполне компьютерное звучание. Опять всем нужен программист... С появлением MPEG-4 наконец-то обретает более реальные и понятные очертания ITV – интерактивное телевидение, о котором давно спорят и под которым каждый понимает что-то свое – от детективов с многовариантным развитием сюжета до простого видео-по-запросу и даже до «ИТВ по-русски» – возможности взять в руки телефон и позвонить в студию.

MPEG-4 появился вовремя, пока еще фактически никто – кроме Apple – не предложил своего готового и работоспособного варианта единой архитектуры доставки мультимедиа, и потому открывается реальная возможность избежать сценария типа «вавилонской башни» в создающейся единой многомиллиардной индустрии. Но ведь для этого комитету MPEG надо было заранее почувствовать тенденцию и начать свою разработку на несколько лет раньше!

...В конце 1996 г. группа MPEG объявила о начале работы над новым стандартом – MPEG-7. В ноябре 1998 г. был закончен прием предложений по этому стандарту, окончательно стандарт принят в сентябре 2001 г. Официально он называется Multimedia Content Description Interface. Чтобы понять, что это такое и почему он необходим, нам придется познакомиться, в дополнение к «контенту», с новым набором понятий цифровой эпохи – метаданные, цифровые медиа-активы, уловить тонкие различия в употреблении английских слов reusing, repurposing, redirection и re-expression (все это применительно к контенту) и, возможно, попытаться подобрать им русские аналоги и так далее...

Литература

1. ISO/IEC JTC1/SC29/WG11 No.2459 Overview of the MPEG-4 Standard. October 1998/Atlantic City.

2. ISO/IEC JTC1/SC29/WG11 No.4668 Overview of the MPEG-4 Standard. March 2002.

3. Koenen R. MPEG-4. Multimedia For Our Time. IEEE Spectrum, February, 1999.

4. Биркмайер К. Общество Плоской Земли. «Мультимедиа. Цифровое видео», №4, 1998.

5. Фоминов О. Мультимедиа и сети. «Мультимедиа. Цифровое видео», №5-6, 7, 8, 1997.